

# THE CAMBRIDGE STATISTICS DISCUSSION GROUP

**Monday 27th January 2014 7:15 for 7:45  
(PLEASE NOTE CHANGE OF DATE TO THAT ORIGINALLY ADVERTISED)**

The Statistical Laboratory,  
Centre for Mathematical Sciences,  
Wilberforce Road, Cambridge, CB3 0WB

## **The UK10K Cohorts Project: Rare variant analysis by whole genome sequencing in 3,621 samples**

**Klaudia Walter**  
The Sanger Institute

**Abstract:** The UK10K Cohorts Project aims to research the relationship between rare and common genetic variants with a comprehensive set of quantitative measures that are relevant to cardiovascular and metabolic disease. The project is based on two cohorts, ALSPAC (Bristol) and TwinsUK (UK-wide), in total 3,621 individuals.

Some of the challenges in this project are to call genetic variants accurately from low-coverage data across thousands of samples and to impute missing and low-quality variants. The loss of power poses another problem for association analyses with rare variants. On the other hand these rare variants allow in depth analyses of the population structure within the UK.

The whole UK10K project received a £10.5 million funding award from the Wellcome Trust, it is currently Britain's largest genomic sequencing consortium (2010 to 2013). The UK10K project also provides an enduring data resource for future research into human genetics.

**Speaker:** Klaudia Walter is currently a statistical geneticist at the Wellcome Trust Sanger Institute. For the past 10 years she has worked on various statistical aspects of genomics, for example on structural variations in the 1000 Genomes Project. She did her PhD with Wally Gilks at the MRC Biostatistics Unit on developing statistical methods for highly conserved non-coding DNA sequences.

**Directions:** The main entrance is reached from Clarkson Road by going along the footpath to the right of the Newton Institute, and turning left through the gatehouse towards the main building (Pavilion A), which has a glass front and a curved grassed roof. The main entrance is in the middle of the glass front. Coffee before the talk will be in the common room in Pavilion D, and the talk will be in Meeting Room 4 in Pavilion A. Free parking is available after 5pm on Clarkson and Wilberforce Roads.

### **Provisional Next Meetings (Please note switch in March and April talks - apologies for any inconvenience):**

20<sup>th</sup> February - Richard Samworth (Statistical Laboratory) on 'High-dimensional variable selection'.

27<sup>th</sup> March - Nick Galwey (GlaxoSmithKline) on 'Working with Epidemiologists'.

30<sup>th</sup> April - Peter Fretwell (British Antarctic Survey).

14<sup>th</sup> October - Adrian Mander (MRC Biostatistics).

17<sup>th</sup> November - Trevor Lewis (TLwise Consulting Limited).

**Supper:** Some members eat before a talk at the University Centre meeting in the Granta Lounge **at 6pm**. All welcome !

**Subscriptions:** of 1 pound are now due for attending the 2013-2014 session.

**Secretary:** Peter Watson, MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 7EF; telephone 01223 355294 Extension 801; E-mail [peter.watson@mrc-cbu.cam.ac.uk](mailto:peter.watson@mrc-cbu.cam.ac.uk)

**Slides and .mp3 files of old talks:** <http://www.mrc-cbu.cam.ac.uk/people/peter.watson/csdg.html>